

Figure 1: The fraction of nodes in the directory that are contained in the largest SCC.

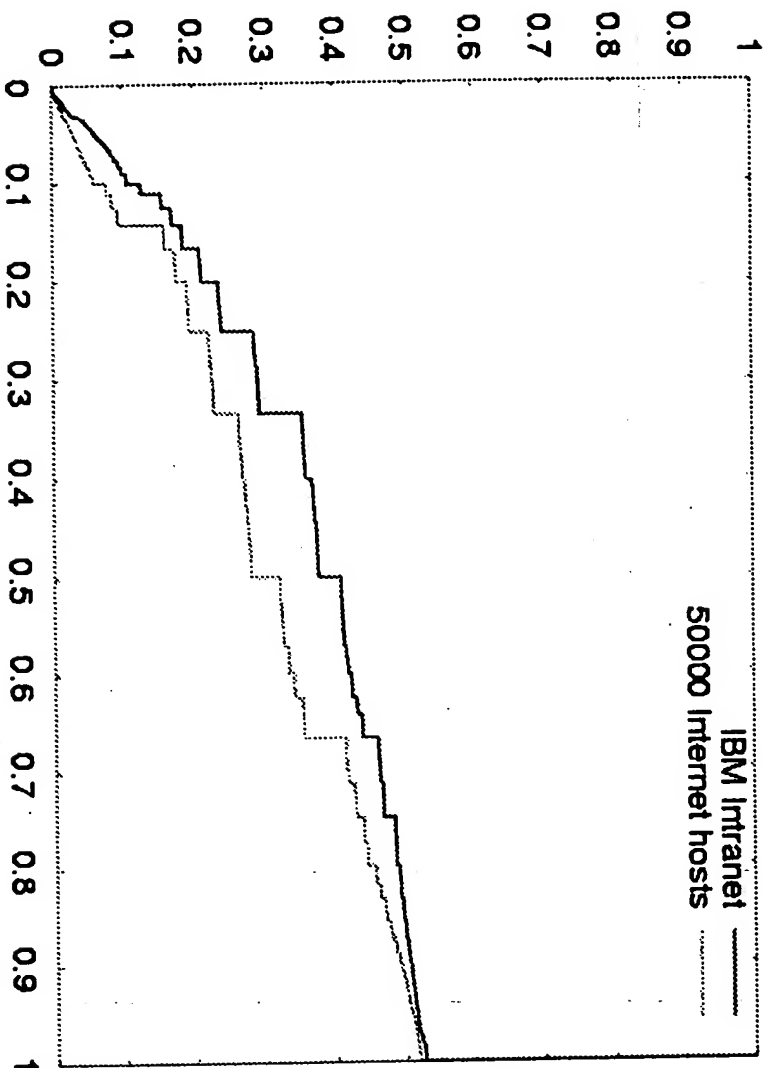


Figure 2: The fraction of nodes in the directory that are contained in the largest reachable component.

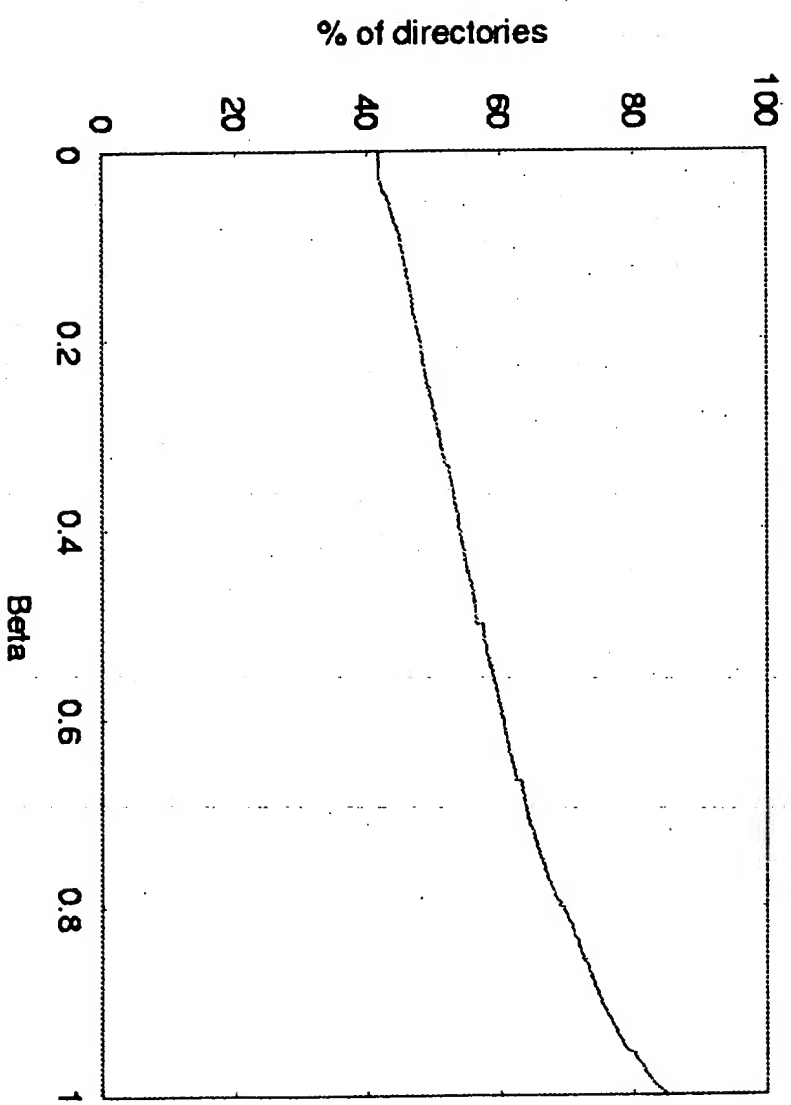


Figure 3: The percentage of directories with a given fraction of common links (for $\alpha = 0.5$).

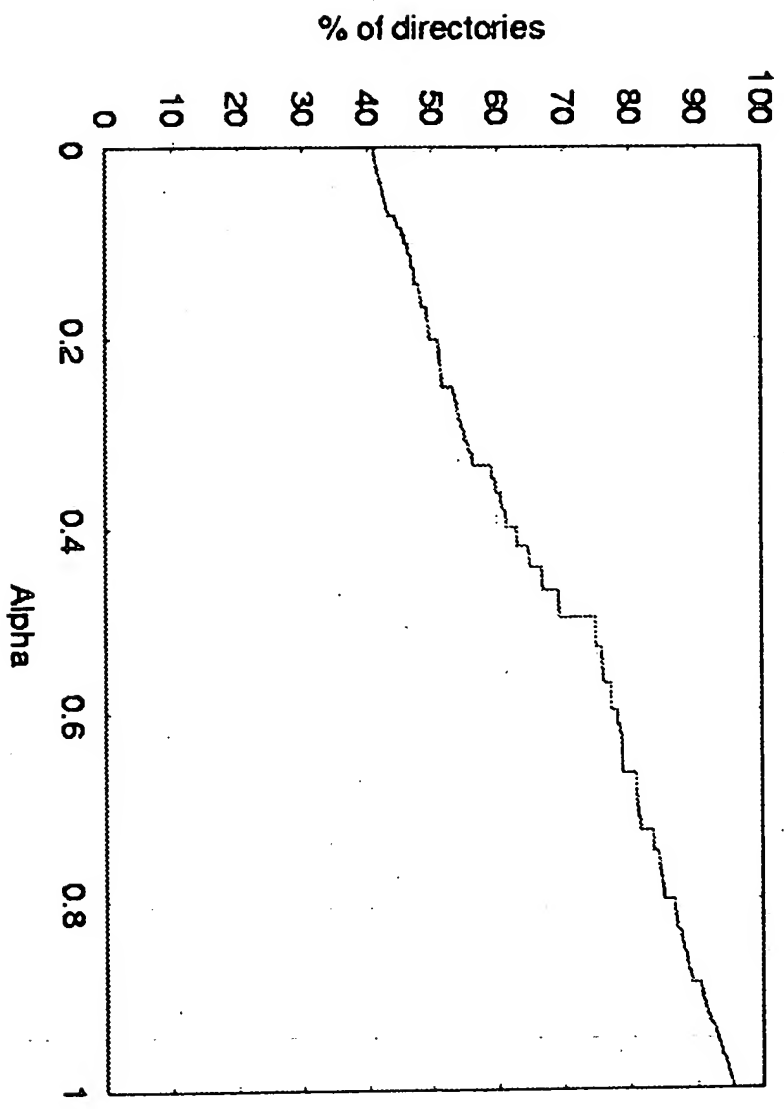


Figure 4: The percentage of directories that have two anchor texts common to an α fraction of the pages.

FIG. 5

